Computers and Electronics in Agriculture 74 (2010) 137–144

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original paper

Hyperspectral waveband selection for internal defect detection of pickling cucumbers and whole pickles

Diwan P. Ariana [a,*], Renfu Lu [b]

[a] Department of Biosystems and Agricultural Engineering, Michigan State University, East Lansing, MI 48824, USA
[b] USDA Agricultural Research Service, Sugarbeet and Bean Research Unit, Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

Article history:
Received 9 March 2010
Received in revised form 13 July 2010
Accepted 28 July 2010

Keywords:
Hyperspectral imaging
Transmittance
Reflectance
Near-infrared
Nondestructive
Waveband selection
Cucumbers
Pickles
Internal defect
Quality

ABSTRACT

Hyperspectral imaging under transmittance mode has shown potential for detecting internal defect, however, the technique still cannot meet the online speed requirement because of the need to acquire and analyze a large amount of image data. This study was carried out to select important wavebands for further development of an online inspection system to detect internal defect in pickling cucumbers and whole pickles. Hyperspectral transmittance/reflectance images were acquired from normal and defective cucumbers and whole pickles using a prototype hyperspectral reflectance (400–740 nm)/transmittance (740–1000 nm) imaging system. Up to four-waveband subsets were determined by a branch and bound algorithm combined with the *k*-nearest neighbor classifier. Different waveband binning operations were also compared to determine the bandwidth requirement for each waveband combination. The highest classification accuracies of 94.7 and 82.9% were achieved using the optimal four-waveband sets of 745, 805, 965, and 985 nm at 20 nm spectral resolution for cucumbers and of 745, 765, 885, and 965 nm at 40 nm spectral resolution for whole pickles, respectively. The selected waveband sets will be useful for online quality detection of pickling cucumbers and pickles.

1. Introduction

Hyperspectral imaging has emerged as a promising technology in recent years for inspecting internal and external quality of foods and agricultural products. It acquires both spatial and spectral information simultaneously from an object, covering a portion of the electromagnetic spectrum, anywhere from ultraviolet to infrared (200 nm to 12 μm), depending on application requirements. A typical configuration of a hyperspectral imaging system consists of lens to gather an image, a dispersive element such as an imaging spectrograph to split the image into spectral channels, and a CCD detector or camera to capture the resultant spatial-spectral images. Properties and characteristics of food and agricultural products vary spatially. Therefore, hyperspectral imaging technique can benefit food quality and safety inspection because of its ability to capture the sample's spectral information in a spatially resolved manner. In comparison, conventional spectroscopic technique cannot provide information on the spatial distribution of chemical compounds in the sample (Peirs et al., 2003). Example applications of hyperspectral imaging include detecting bruises and other defects on apples (Lu, 2003; Mehl et al., 2004; Nicolaï et al., 2006; Xing et al., 2007), fecal contamination on apples and poultry carcasses (Kim et al., 2002; Park et al., 2005), quality attributes on strawberry (El Masry et al., 2007), firmness and soluble solids content of apples (Lu, 2007; Peng and Lu, 2008), ripeness on tomatoes (Polder et al., 2002), canker on citrus (Qin et al., 2009), freeze damage in mushroom (Gowen et al., 2009), and bruise and internal defect on pickling cucumbers and pickles (Ariana and Lu, 2008b, 2010; Ariana et al., 2006).

Hyperspectral images are three-dimensional data containing two spatial dimensions and one spectral dimension. A hyperspectral imaging system with high spectral and spatial resolutions will generate a large amount of data. It is thus challenging to implement the technology for high-speed online inspection of food and agricultural products. One way of overcoming this problem is to implement hyperspectral imaging technology in multispectral imaging mode (Chao et al., 2010). This would require identification of a few most useful wavelengths. Another downside of the technique is that the images at different spectral bands are strongly correlated (De Backer et al., 2005). Highly redundant data may cause convergence instability in the development of prediction or classification models (Bajcsy and Groves, 2004). A good practice of accurately predicting an unknown test sample is to use at least 10 times as many training samples per class as the number of features (i.e., number of wavebands in the hyperspectral imaging case). This

* Corresponding author at: 105AA Farrall Hall, Michigan State University, East Lansing, MI 48824, USA. Tel.: +1 517 432 7438; fax: +1 517 432 2892.
E-mail address: arianadi@msu.edu (D.P. Ariana).

0168-1699/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.compag.2010.07.008

is known as the curse of dimensionality. If too many features are used and the available training set is too small, poor generalization is expected (Jain et al., 2000). In many cases, a large number of training samples are difficult to acquire, therefore reducing the number of features in a model is required, either by feature selection or feature extraction technique. A feature selection algorithm selects the best small subset from the original feature set, while a feature extraction algorithm creates new features based on transformations of the original feature set. Feature selection technique is a preferred choice for applications that require fast data acquisition such as in an online inspection system. Thus, hyperspectral imaging technique is frequently used as a research tool to identify important wavebands that can be implemented in an online multispectral (using only a few discrete wavebands) machine vision system.

Currently, machine vision systems are commonly used in many horticultural product packing or processing facilities for inspection of external features (i.e., size, shape, color, and/or surface defects) but not for internal defects. Pickling cucumbers are prone to internal defect, which often occurs in the form of carpel separation or hollow centers, caused by undesirable growing conditions or excessive mechanical load during harvest, transport and postharvest handling (Miller et al., 1995). These defective cucumbers, if not segregated, would cause a bloating problem during brining. At present, sorting for internal defect is only carried out on whole desalted pickles through manual visual inspection and hand touch, which is unreliable and labor intensive (Ariana and Lu, 2008a). It is therefore desirable that a machine vision system be used for removing defective pickling cucumbers and pickles from normal ones prior to and also after brining. Hyperspectral imaging under transmittance mode has been studied for detection of internal defect in fresh cucumbers and pickles (Ariana and Lu, 2008b, 2010). However, speed and cost are the main constraining factor for implementation of the technique in commercial processing lines for evaluation of individual cucumbers and pickles. Therefore, this study was focused on determining the optimal wavebands that are useful for online detection of internal defect in pickling cucumbers and whole pickles. This would help to assure the quality and consistency of final pickled products, which are estimated to be worth billions of dollars for the U.S. pickle industry (Lucier and Jerardo, 2007).

2. Materials and methods

2.1. Pickling cucumber and whole pickle samples

'Journey' is one of the most commonly grown cucumber variety in USA. During the harvest session in 2008, 'Journey' pickling cucumbers were hand harvested at an optimal maturity stage and were carefully handled after harvest to ensure no mechanical injury to the samples from the experimental fields of Michigan State University in Oceana County, Michigan. The cucumbers were washed, and only the ones that were free from external defect and/or misshapen and with the diameters of 37–55 mm (commercial size 3) were selected for the experiment. Prior to the sample selection, a few samples of harvested cucumbers were sliced for visual inspection to ensure that the cucumbers were free from internal defect that may have been caused by adverse growing conditions and/or improper handling. None of the fresh cucumbers showed evidence of internal defects. Hence it was reasonable to assume that the rest of the samples were free from internal defect. A total of 300 fresh cucumber samples were then selected to be used in the experiment.

In a separate experiment, desalted whole pickles of medium size (35–38 mm in diameter) were collected from a commercial pickle processing facility in Pinconning, Michigan, USA in September 2008. The information about the cultivars of the samples was not available from the pickle processor because different cultivars were mixed during the process. The pickle samples were collected from two pickle processing lanes after manual sorting for bloating damage; one lane carried rejected whole pickles for relish and the other lane carried good pickles. One hundred forty pickles each of good and defective classes were randomly selected for imaging.

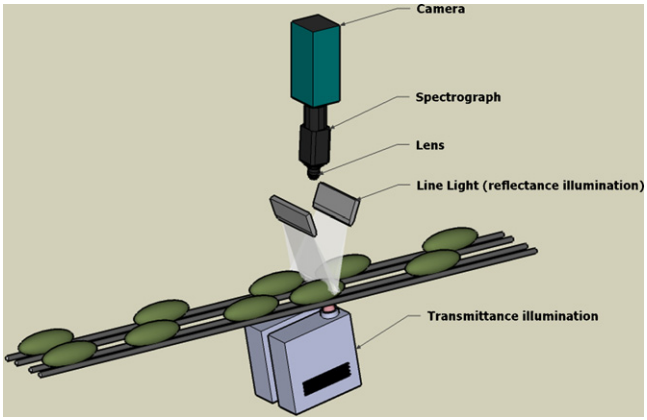

Fig. 1. Schematic of the hyperspectral imaging system (Ariana and Lu, 2010).

2.2. Hyperspectral imaging system

A prototype hyperspectral imaging system developed in the USDA/ARS postharvest engineering laboratory at Michigan State University, East Lansing, Michigan was used for this research (Ariana and Lu, 2010). The schematic of the prototype, which consists of imaging, illumination and conveying systems, is shown in Fig. 1. The prototype was designed for evaluation of external and internal quality of pickling cucumber using simultaneous reflectance and transmittance imaging modes. However, for this study, only internal quality was evaluated. The imaging system consisted of a 1376 × 1040 pixel CCD camera (Sensicam QE, The Cooke Corp., Romulus, MI, USA), an imaging spectrograph (ImSpector V10E OEM, Spectral Imaging Ltd., Oulu, Finland) covering the spectral range of 400–1000 nm, and a 23 mm fixed focal lens (Xenoplan, Schneider Optics, Hauppauge, NY, USA). Two illumination sources were used in the prototype; one was for reflectance and the other for transmittance mode. A 150 W tungsten halogen lamp with a DC-regulated power supply (Fiberlite A240P, Dolan-Jenner Industries, Lawrence, MA, USA) connected to a dual fiber optic line light was used for visible reflectance illumination for 400–740 nm by means of a short-pass filter placed between the lamp and the fiber optics. This avoided the mixing of reflectance with the transmittance signal in the near-infrared region from the second light source, which consisted of a 410 W tungsten halogen lamp with a built-in reflector (FXL/5, Eiko Ltd., Shawnee, KS, USA). The lamp was housed in a custom made enclosure containing mirror and lens to direct the light to the samples moving on the belts. One light enclosure was installed underneath the top portion of the belts for each lane. Extruded aluminum framing and urethane belts with the diameter of 13 mm were used to construct a 1.70 m long two-lane round-belt conveyor. Each lane was constructed with two urethane belts driven by pulleys. The two belts, separated by 25 mm distance, were designed to transport commercial size three pickling cucumbers (37–55 mm diameter). However, the distance between the two belts could be changed by adjusting the distance between pulleys in their common shafts to accommodate different sizes of cucumbers. Using two round-belts in the conveyor allows imaging in transmittance mode by placing the light source underneath the top part of the belts. The transmittance light reached the lower part of the sample in between the two belts, where cucumbers were transported.

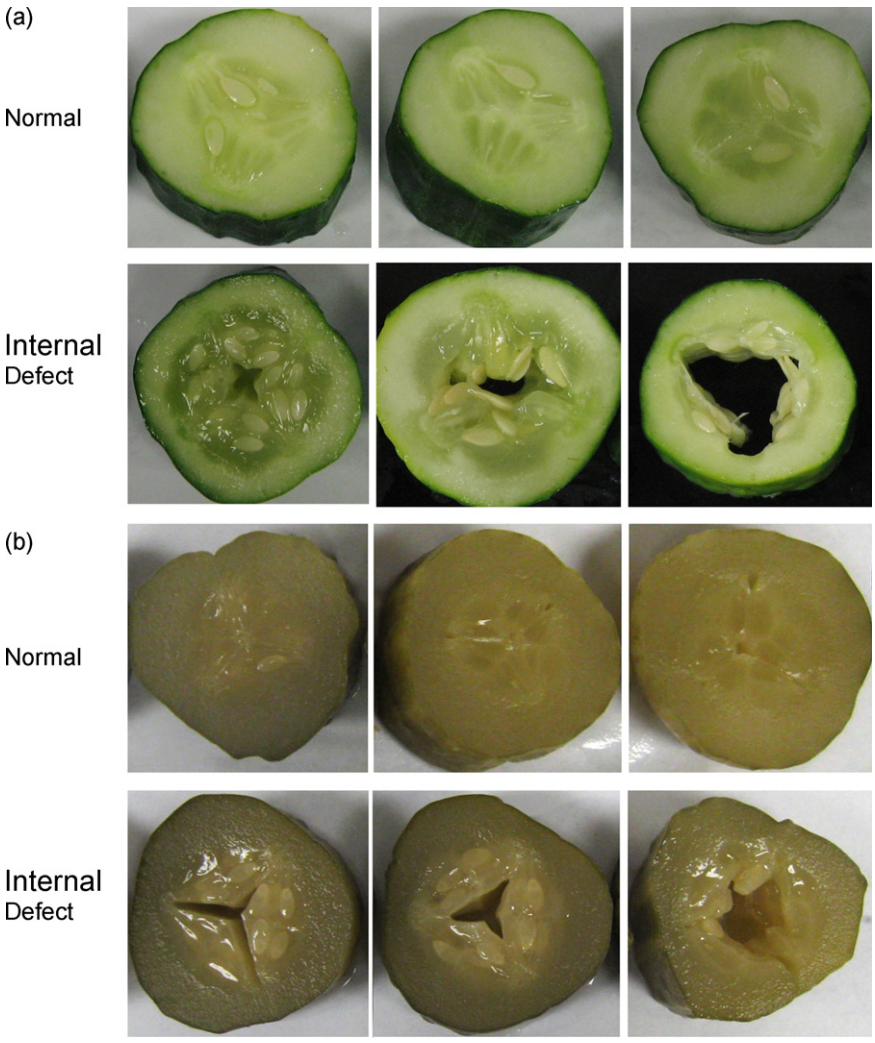

Fig. 2. Normal and defective slices of (a) fresh pickling cucumbers and (b) pickles.

A more detailed description of the prototype is given in Ariana and Lu (2010).

2.3. Image acquisition and pre-processing

The camera was set to run under continuous mode with 2 ms exposure time and 8×8 binning in spectral and spatial dimensions, which gave a spectral resolution of 5 nm, approximately the same as the spectral resolution of the imaging spectrograph and 1 mm spatial resolution. Samples were transported to the field of view of the imaging system with a conveyor belt running at the speed of 110 mm/s or approximately up to two cucumbers per second. The first set of hyperspectral images were acquired from the 300 cucumbers before they were subjected to mechanical stress to induce internal damage. These images represented normal cucumbers, providing the ground truth information. Each cucumber was then subjected to rolling between a bench top and a 30 cm square board supporting a 10 kg load. Two different rolling times (15 and 30 s) were applied to two groups of cucumbers (150 each) in order to introduce light and severe internal damage. The second set of hyperspectral images were acquired from the mechanical stress-treated cucumbers, which represented the internal defect class. For the second experiment, hyperspectral images were acquired from the 140 desalted whole pickle samples collected from the good and bloated lanes, respectively.

To build a supervised classification model, a class has to be assigned to each of the samples. After imaging, the cucumbers and whole pickles were sliced diagonally into six pieces of approximately the same thickness for visual inspection of their internal quality. Grade 0 or 1 was assigned to each slice for normal and damaged endocarp (seed cavity), respectively. A normal slice (grade 0) was characterized by seeds being intact within the seed cavity, whereas a defective slice (grade 1) was characterized by the formation of water-soaked lesions in the endocarp (seed cavity) and/or hollow/split at the center (Fig. 2).

Only the wavebands between 500 and 1000 nm were used for data analysis and algorithm development, because the signals beyond this spectral region were too low and noisy to be useful. At 5 nm spectral resolution, the number of wavebands for analysis was 101. Spectral calibration of the samples was carried out by using information obtained from the in-line calibration materials placed in the field of view of the imaging system. The intensity values for the visible range from 500 to 740 nm were calibrated by a reflectance reference, made from PVC, which was placed in between the lanes, while the intensity values in the spectral range of 740–1000 nm were calibrated by two transmittance references, made from white Teflon, placed at the outer side of each lane. A more detailed explanation on the in-line light calibration of the prototype hyperspectral imaging system can be found elsewhere (Ariana and Lu, 2008b).

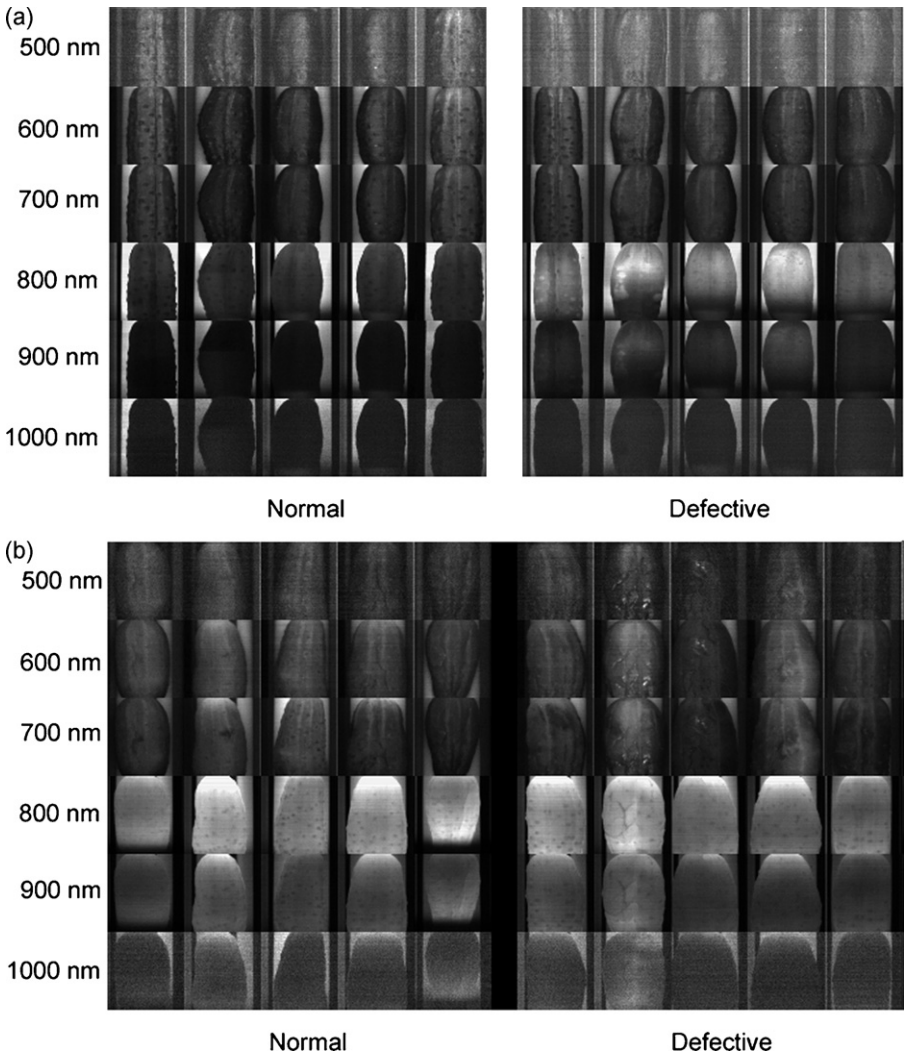

Fig. 3. Spectral images for the normal (left) and defective (right) groups of (a) fresh pickling cucumbers and (b) pickles.

2.4. Wavebands selection

Hyperspectral images, as high-dimensional data, exhibit a high degree of interband correlation, leading to data redundancy that can cause convergence instability in classification models. Therefore, the use of fewer wavebands is preferable for more stable classification and easier implementation in a multispectral imaging system to meet the speed requirement of a sorting line. A branch and bound algorithm was used in this study to select up to four wavebands, suitable for implementation in a multispectral imaging system. Furthermore, four software binning operations (i.e., 1×1, 1×2, 1×4, 1×8, and 1×12) were performed, resulting in the spectral resolutions of 5, 10, 20, 40, and 60 nm for the resultant hyperspectral images with the spatial resolution remaining unchanged. The binning operations would help to select an appropriate bandwidth for the bandpass filters to be used in a multispectral imaging system.

For a feature selection problem, one tries to find the best set of d features from $\binom{D}{d} = \frac{D!}{d!(D-d)!}$ possible sets, where d is the number of selected features and D is the total number of features. Each set is evaluated with a selection criterion, and only the one which results in the highest classification accuracy is chosen. In practice, however, this often is not feasible because $\binom{D}{d}$ is very large and computationally excessive for sets of even moderate size. However, it may be possible to determine the optimal feature set without explicit evaluation of all the possible combinations of d measurements with the help of the branch and bound algorithm. The algorithm is applicable under the assumption that a feature selection criterion satisfies the monotonicity property, i.e., the performance of a feature subset should improve whenever a feature is added to it (Jain et al., 2000). The algorithm starts searching from the original feature set and it proceeds by removing features from it. A bound is placed on the value of the evaluation function to create a rapid search. As the evaluation function obeys the monotonicity principle, any subset for which the value is less than the bound is removed from the search tree (i.e., all subsets of it are discarded from the search space). Examples of commonly used evaluation functions are Mahalanobis distance, discriminant function, Fischer criterion, and Bhattacharya distance (Dash and Liu, 1997).

The fresh cucumber and whole pickle samples were split into training and validation sets with the ratio of 75 and 25%, respectively. Pixels from the masked samples training set were used for wavebands selection using the branch and bound algorithm with Mahalanobis distance as the criterion function since it is monotonic. Finally, using the selected wavebands, k-nearest neighbor (KNN) classifier with Euclidean distance measure and $k = 1$ was used for pixel classification in the validation set. Images that contained more than 5% defective pixels were classified as defective; otherwise,

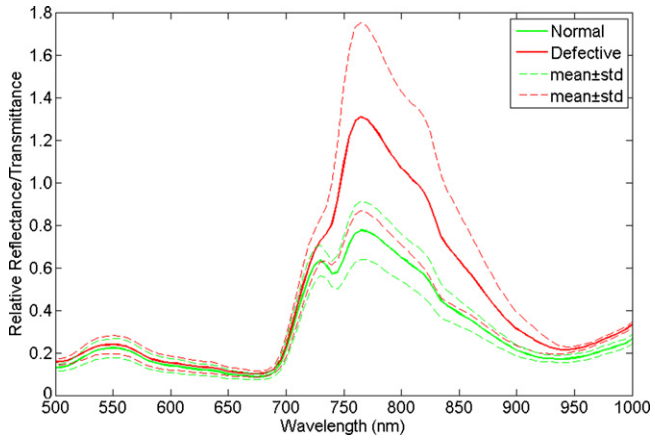

Fig. 4. Mean reflectance (500–740 nm) and transmittance (740–1000 nm) spectra of normal and defective cucumbers.

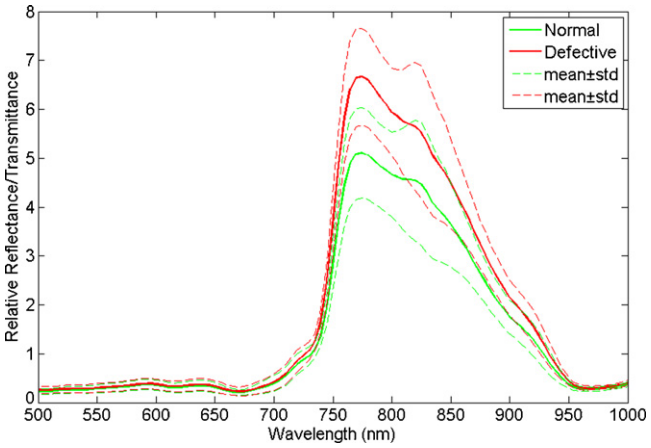

Fig. 5. Mean reflectance (500–740 nm) and transmittance (740–1000 nm) spectra of normal and defective whole pickles.

they would be classified as good. The classification performance of the feature selection method was also compared to the feature extraction method using partial least square discriminant analysis (PLSDA). MATLAB programs using the Image Processing and Statistics Toolboxes (The MathWorks, Inc., Natick, MA, USA) and PLS Toolbox (Eigenvector Research, Inc., Wenatchee, WA, USA) were written for performing spectra extraction from the hyperspectral images, waveband selection and classification.

3. Results and discussion

3.1. Spectral characteristics

The hyperspectral imaging system employed the simultaneous reflectance and transmittance imaging modes. The spectral images below 740 nm were reflectance images, which carried mostly color or surface features information; meanwhile the spectral images above 740 nm were transmittance images, containing internal characteristics of the cucumber or pickle. Representative hyperspectral images of normal and defective pickling cucumbers and whole pickles at six selected wavebands from 500 to 1000 nm are shown in Fig. 3. The pixel values at some wavebands were low, resulting in dark images for both normal and defective pickling cucumbers or whole pickles, which were hardly distinguished. When the images of normal and defective samples at the same waveband were scaled to maximum contrast using the same factor, their differences were enhanced, as demonstrated in Fig. 3a. Images at 800 and 900 nm appeared brighter for the defective cucumbers than for the normal cucumbers. This difference was not apparent in whole pickles (Fig. 3b). In some severely defective cucumbers, the bright areas appeared more intense compared to the surrounding pixels on the images at 800 nm (e.g., the second cucumber from the left for the defective group in Fig. 3a). These bright areas did not appear in the visible range. The light scattering abilities of cellular components (e.g., cell walls and starch granules), which normally diffract or reflect light, might have decreased due to the fluid build-up from the ruptured cells (Miller et al., 1995). Severely defective cucumbers would transmit more light because they had a substantial portion of the endocarp missing, which was consequently filled with air.

The mean spectra and associated standard deviations of the normal and defective cucumbers and whole pickles are shown in Figs. 4 and 5, respectively. The relative transmittance of whole pickles was much higher (approximately six times) compared to that of fresh cucumbers. Differences in the tissue texture between fresh cucumbers and pickles are a probable cause for their difference in transmittance. Since pickles have more homogenous and watery texture than cucumbers, they have lesser interfaces for diffracting light, which could translate into more transmittance. In general, signal was stronger in the NIR region than in the visible region. Both normal and defective cucumber spectra exhibited strong absorption at 680 and 950 nm due to chlorophylls and water, respectively. The spectra of the cucumbers showed local maxima at 550 nm, which represented the dark green color of the cucumber surface. The local maxima did not appear on the olive-green colored pickles. Transmittance at 775 nm was stronger for defective than normal cucumbers or whole pickles, which is consistent with the images shown in Fig. 3. More overlapping in the spectra of normal and defective whole pickles was observed than in the spectra of pickling cucumbers, especially above 800 nm (Fig. 4 vs. Fig. 5). This suggested the difficulty in classifying the whole pickles. The overlapping spectra in the visible region of 500–725 nm for the normal and defective cucumbers or whole pickles suggested that both classes had similar color and that internal defect was not visible from the surface. Thus it would be difficult to segregate defective cucumbers/whole pickles from normal ones using hyperspectral reflectance images. Rolling treatments of 15 and 30 s to the cucumbers did not result in significantly different levels of damage, as shown in Fig. 6, and also from visual observation after the cucumbers were sliced.

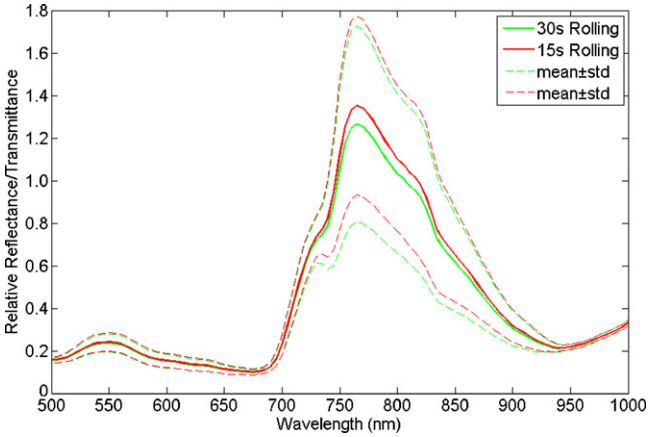

Fig. 6. Mean reflectance (500–740 nm) and transmittance (740–1000 nm) spectra of defective cucumbers after 10 kg load applied for 15 and 30 s.

Table 1
Classification accuracies of pickling cucumbers based on the selected wavebands.

Spectral resolution	Number of bands	Center wavelengths	Normal	Defective	Total
5 nm	2	900 965	86.8	81.8	85.3
	3	975 900 965	92.5	90.9	92.0
	4	840 975 900 965	94.3	90.9	93.3
10 nm	2	975 965	86.8	81.8	85.3
	3	805 975 965	92.5	90.9	92.0
	4	955 805 975 965	92.5	95.5	93.3
20 nm	2	545 965	88.7	86.4	88.0
	3	805 985 965	90.6	95.5	92.0
	4	745 805 985 965	94.3	95.5	94.7
40 nm	2	805 965	90.6	81.8	88.0
	3	885 925 965	92.5	95.5	93.3
	4	845 725 925 965	96.2	90.9	94.5
60 nm	2	845 965	86.8	81.8	85.3
	3	905 785 965	92.5	90.9	92.0
	4	725 905 785 965	94.3	90.9	93.3

3.2. Wavebands selection

Classification results for two, three, and four selected wavebands for different spectral resolutions are summarized in Tables 1 and 2 for the validation sets of cucumbers and whole pickles respectively. In general, classification accuracies increased as more wavelengths were selected; they were up to 88.0, 93.3, and 94.7% for pickling cucumbers and up to 77.1, 80.0, and 82.9% for whole pickles for the 2-, 3-, and 4-waveband models, respectively. The highest total classification accuracy for cucumbers was 94.7% at the spectral resolution of 20 nm with the four selected wavebands of 745, 805, 965, and 985 nm, while the highest total classification accuracy for whole pickles was 82.9% at the 40 nm spectral resolution with the selected wavebands of 745, 765, 885, and 965 nm. Classification accuracies for whole pickles were lower, which would have been expected in view of the spectral characteristics of normal and defective whole pickles (Figs. 4 and 5). Most of the selected wavebands were in the near-infrared region, which is explicable from the spectral characteristics and also in agreement with a previous study (Ariana and Lu, 2008b).

When the spectral resolution changed from 5 to 10 nm, the classification results for cucumbers were the same for all three waveband combinations (Table 1). However, as the spectral resolution changed to 20 nm, the total classification accuracies improved from 85.3 to 88.0% and from 93.3 to 94.7% for the two- and four-waveband models, respectively. A further decrease in the spectral resolution to 40 nm did not improve classification accuracy. However, the classification accuracies decreased at 60 nm resolution, and they might further decline as the spectral resolution became lower. These results suggested that a spectral resolution between 20 and 40 nm would be appropriate for defect detection of pickling cucumbers.

For whole pickles, the spectral resolution of 40 nm had better results for all three waveband combinations, compared to 5, 10, 20 and 60 nm resolutions. This suggested that high spectral resolution (i.e., 5 nm) is not needed for defect detection of cucumbers or pickles, but excessively low spectral resolutions (e.g., 60 nm) would also be detrimental to classification accuracy. Hence, we can use larger hardware binning in the spectral dimension during image acquisition, while maintaining the same spatial resolution, to improve the image acquisition and processing speed. Spectral binning reduces spectral resolution, but improves the image quality by increasing the signal-to-noise ratio (Lu, 2003). This finding further suggests that in designing a multispectral imaging system using bandpass filters, larger bandwidth filters are desirable because they can pass more signal to the CCD sensor.

It should be mentioned that some of the selected wavebands are in the spectral region of 900–1000 nm, over which the camera had a relatively low signal-to-noise ratio. However, this spectral range was still considered in the analysis because the degree of spectral separation between normal and defective fresh cucumbers for this spectral region was quite high (Fig. 4). Hence in future implementation of these wavebands for online

Table 2
Classification accuracies of whole pickles based on the selected wavebands.

Spectral resolution	Number of bands	Center wavelengths	Normal	Defective	Total
5 nm	2	820 935	61.5	68.2	65.7
	3	820 715 935	69.2	68.2	68.6
	4	785 730 715 935	61.5	72.7	68.6
10 nm	2	560 935	76.9	63.6	68.6
	3	780 715 935	76.9	72.7	74.3
	4	780 695 715 935	69.2	81.8	77.1
20 nm	2	905 785	69.2	72.7	71.4
	3	865 745 905	84.6	72.7	77.1
	4	725 865 905 785	84.6	77.3	80.0
40 nm	2	645 885	84.6	72.7	77.1
	3	745 965 885	92.3	72.7	80.0
	4	765 745 965 885	92.3	77.3	82.9
60 nm	2	665 905	76.9	63.6	68.6
	3	785 965 905	69.2	72.7	71.4
	4	605 785 965 905	84.6	72.7	77.1

Table 3
Classification accuracies of pickling cucumbers based on partial least square discriminant analysis (PLSDA) and *k*-nearest neighbor (KNN) using all wavebands.

Spectral resolution	Classifiers	Normal	Defective	Total
5 nm	PLSDA	96.2	95.5	96.0
	KNN	89.5	80.0	86.7
10 nm	PLSDA	98.1	95.5	97.3
	KNN	90.5	82.2	88.0
20 nm	PLSDA	96.2	95.5	96.0
	KNN	89.5	80.0	86.7
40 nm	PLSDA	94.3	95.5	94.7
	KNN	87.6	77.8	84.7
60 nm	PLSDA	93.3	91.1	92.7
	KNN	85.7	75.6	82.7

sorting and grading, it is important to select an appropriate camera with an adequate spectral response for the wavelengths of 900–1000 nm.

In addition, classification based on the entire set of wavebands from 500 to 1000 nm using partial least square discriminant analysis (PLSDA) and KNN was also performed to compare with the selected waveband approach (Tables 3 and 4). The highest classification accuracies of 97.3 and 88.0% for cucumbers were obtained for PLSDA and KNN, respectively, at 10 nm spectral resolution (Table 3). For whole pickles, the highest classification accuracies were 88.0 and 70.0% for PLSDA and KNN, respectively, at 20 nm spectral resolution (Table 4). Overall the PLSDA performed better than KNN using all wavebands. The results from PLSDA using all wavebands were also better than those from KNN using a few selected wavebands (Tables 1 and 2). However, the results from KNN using all wavebands were generally not as good as those obtained using three or four selected wavebands. These results demonstrated the effectiveness of PLSDA as a feature extraction method. Although the classification results with three or four wavebands were not as good as those with the full wavelength range using PLSDA, the difference for the two methods at their respective optimal spectral resolutions was only 2.6 percentage points. Hence it is preferred to use fewer wavelengths for online implementation because of faster image acquisition and processing.

Representative images for the classification results for cucumbers are shown in Fig. 7. Coloration has been added to the cucumbers and defective areas to enhance visual observation. Most of the defective areas in the segmented images appeared large, although smaller areas were also identified, which might have been due to damage in the form of water-soaked lesion in the mesocarp region near the surface.

Table 4
Classification accuracies of whole pickles based on partial least square discriminant analysis (PLSDA) and *k*-nearest neighbor (KNN) using all wavebands.

Spectral resolution	Classifiers	Normal	Defective	Total
5 nm	PLSDA	88.7	77.3	85.3
	KNN	65.8	62.5	64.3
10 nm	PLSDA	88.7	81.8	86.7
	KNN	68.4	68.8	68.6
20 nm	PLSDA	90.6	81.8	88.0
	KNN	71.1	68.8	70.0
40 nm	PLSDA	90.6	77.3	86.7
	KNN	68.4	65.6	67.1
60 nm	PLSDA	89.5	75.0	82.9
	KNN	65.8	62.5	64.3

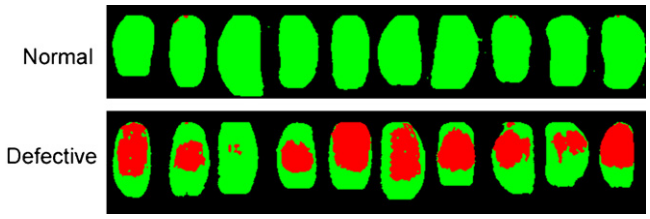

Fig. 7. Segmented images of normal and defective cucumbers (damage areas are denoted with red color). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

4. Conclusions

Transmittance mode in the near-infrared region of 740–1000 nm was effective for detection of internal defect in cucumbers and pickles. Transmittance for internally defective cucumbers or pickles was generally higher than that for normal ones. Wavebands at 745, 805, 965, and 985 nm with 20 nm spectral resolution were the best four-waveband set for cucumbers, achieving an overall classification accuracy of 94.7%. For whole pickles, the four-waveband set at 745, 765, 885, and 965 nm with 40 nm spectral resolution yielded the best classification accuracy of 82.9%. This research demonstrated that high spectral resolution is not necessary when using the selected wavebands for internal defect detection of cucumbers and pickles. Finally, an imaging sensor with good quantum efficiency in the spectral region of 900–1000 nm is needed for implementing the selected wavelengths for online sorting and grading of cucumbers.

Acknowledgment

The authors wish to thank Mr. Benjamin Bailey, Engineering Technician, for providing technical support to this research.

References

Ariana, D.P., Lu, R., 2008a. Detection of internal defect in pickling cucumbers using hyperspectral transmittance imaging. Trans. ASABE 51, 705–713.
Ariana, D.P., Lu, R., 2008b. Quality evaluation of pickling cucumbers using hyperspectral reflectance and transmittance imaging. Part II. Performance of a prototype. Sens. Instrum. Food Qual. Saf. 2, 152–160.
Ariana, D.P., Lu, R., 2010. Evaluation of internal defect and surface color of whole pickles using hyperspectral imaging. J. Food Eng. 96, 583–590.
Ariana, D.P., Lu, R.F., Guyer, D.E., 2006. Near-infrared hyperspectral. reflectance imaging for detection of bruises on pickling cucumbers. Comput. Electron. Agric. 53, 60–70.
Bajcsy, P., Groves, P., 2004. Methodology for hyperspectral band selection. Photogramm. Eng. Remote Sensing 70, 793–802.
Chao, K., Yang, C.-C., Kim, M.S., 2010. Spectral line-scan imaging system for high-speed non-destructive wholesomeness inspection of broilers. Trends in Food Science and Technology 21, 129–137.
Dash, M., Liu, H., 1997. Feature selection for classification. Intell. Data Anal. 1, 131–156.
De Backer, S., Kempeneers, P., Debruyn, W., Scheunders, P., 2005. A band selection technique for spectral classification. IEEE Geosci. Remote Sensing Lett. 2, 319–323.
El Masry, G., Wang, N., El Sayed, A., Ngadi, M., 2007. Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry. J. Food Eng. 81, 98–107.
Gowen, A.A., Taghizadeh, M., O'Donnell, C.P., 2009. Identification of mushrooms subjected to freeze damage using hyperspectral imaging. J. Food Eng. 93, 7–12.
Jain, A.K., Duin, R.P.W., Mao, J., 2000. Statistical pattern recognition: a review. IEEE Trans. Pattern Anal. Mach. Intell. 22, 4–37.
Kim, M.S., Lefcourt, A.M., Chao, K., Chen, Y.R., Kim, I., Chan, D.E., 2002. Multispectral detection of fecal contamination on apples based on hyperspectral imagery. Part I. Application of visible and near-infrared reflectance imaging. Trans. ASAE 45, 2027–2037.
Lu, R., 2003. Detection of bruises on apples using near-infrared hyperspectral imaging. Trans. ASAE 46, 523–530.
Lu, R., 2007. Nondestructive measurement of firmness and soluble solids content for apple fruit using hyperspectral scattering images. Sens. Instrum. Food Qual. Saf. 1, 19–27.

Lucier, G., Jerardo, A., 2007. Commodity Highlight: Pickling Cucumbers, Vegetables and Melons Outlook. Economic Research Service, United States Department of Agriculture.

Mehl, P.M., Chen, Y.R., Kim, M.S., Chan, D.E., 2004. Development of hyperspectral imaging technique for detection of apple surface defects and contaminations. J. Food Eng. 61, 67–81.

Miller, A.R., Kelley, T.J., White, B.D., 1995. Nondestructive evaluation of pickling cucumbers using visible-infrared light transmission. J. Am. Soc. Hortic. Sci. 120, 1063–1068.

Nicolaï, B.M., Lötze, E., Peirs, A., Scheerlinck, N., Theron, K.I., 2006. Non-destructive measurement of bitter pit in apple fruit using NIR hyperspectral imaging. Postharvest Biol. Technol. 40, 1–6.

Park, B., Lawrence, K.C., Windham, W.R., Smith, D.P., 2005. Detection of cecal contaminants in visceral cavity of broiler carcasses using hyperspectral imaging. Appl. Eng. Agric. 21, 627–635.

Peirs, A., Scheerlinck, N., Baerdemaeker, J.D., Nicolai, B.M., 2003. Starch index determination of apple fruit by means of a hyperspectral near infrared reflectance imaging system. J. Near Infrared Spec. 11, 379–389.

Peng, Y., Lu, R., 2008. Analysis of spatially resolved hyperspectral scattering images for assessing apple fruit firmness and soluble solids content. Postharvest Biol. Technol. 48, 52–62.

Polder, G., Heijden, G.W.A.M., Young, I.T., 2002. Spectral image analysis for measuring ripeness of tomatoes. Trans. ASAE 45, 1155–1161.

Qin, J., Burks, T.F., Ritenour, M.A., Bonn, W.G., 2009. Detection of citrus canker using hyperspectral reflectance imaging with spectral information divergence. J. Food Eng. 93, 183–191.

Xing, J., Saeys, W., De Baerdemaeker, J., 2007. Combination of chemometric tools and image processing for bruise detection on apples. Comput. Electron. Agric. 56, 1–13.